

REPLACED BY
ART 34 AMDT

10/5/05 1229

is constrained to numbers only, further restricting the potential recognition alternatives.

Zip/Postal Code

Zip/Postal codes within a given country generally follow a specific pattern. For example: in Australia, the postal codes are always four digits long; in the USA, five digits; and in the UK, a mix of one or more letters, followed by two or more numbers, followed by one or more letters again. Additional decoding constraints are available if the corresponding State and Suburb information is available.

Country, Region, etc.

Full lists of possible Country/Region labels are publicly available.

Birth Date, Date of Birth, Other dates etc.

Written dates generally follow a regular pattern, and have a constrained character set consisting of either numbers alone or numbers and delimiting characters such as '-' or '/'.

Email, E-Mail, Email Address, etc.

Email addresses follow a specific pattern and have a well-specified character set. An example regular expression that can be used to match email addresses is `"/^([a-zA-Z0-9_\\.\\-])+\\@((([a-zA-Z0-9\\-])+\\.)+([a-zA-Z0-9])+$/"`.

In addition to this, if email contact information is available for a user (e.g. using Microsoft Windows Messaging API (MAPI)), the list of email addresses can be used as a dictionary during recognition. Similarly, common email domain names (e.g. "hotmail.com", "yahoo.com", "email.com", etc.) can be used as dictionary entries to guide recognition.

Credit Card, Credit Card Number, etc.

Credit card numbers have a specific format (e.g. "####-####-####-####") and constrained character set.

What is claimed is:

1. A method of interpreting data input to a form-based data entry system, including decoding data entered into a particular form field such that its information content can be
5 determined, said information content being in a consistent machine-readable format, wherein said decoding of data includes determining one or more possible values of information content, certain pre-defined possible outcomes being given a relatively higher probability of being correct, and said pre-defined possible outcomes being dependent on the context of the particular form field.
- 10 2. The method as claimed in claim 1, wherein said decoding of data is performed contemporaneously with the data entry (online).
3. The method as claimed in claim 1, wherein said decoding of data is performed
15 some time after the data entry (offline).
4. The method as claimed in any one of the preceding claims, wherein the data entry is effected by one or both of handwritten characters and speech.
- 20 5. The method as claimed in any one of the preceding claims, wherein the particular form field has associated with it a pre-defined dictionary of possible decoded data, said dictionary being used to constrain the decoding process.
6. The method as claimed in claim 5 wherein, certain entries in the dictionary are
25 assigned a higher probability of being the correct decoded data.
7. The method as claimed in either of claims 5 or 6, wherein the field is a name field and the predefined dictionary includes an indication of gender associated with selected names.
- 30 8. The method as claimed in either of claims 5 or 6, wherein the field is an address field having sub-fields arranged hierarchically such that a decoded entry in a sub-field may be used to constrain an entry in another sub-field.

9. The method as claimed in either of claims 5 or 6, wherein the field is a telephone number field and is constrained such that the only valid data includes numerals only.
10. The method as claimed in any one of the preceding claims, wherein the field is a credit card number, wherein the only valid data includes a fixed number of numerals, said numerals being further verifiably by use of a checksum.
11. The method as claimed in any one of the preceding claims, wherein the field is from the set including: zip/post code; country; date; email address; and/or language.
12. The method as claimed in any one of the preceding claims, wherein the said system is implemented using one of the standardized file formats: HTML, XML, PDF and XForms.
13. The method as claimed in any one of the preceding claims, wherein a custom validation program is associated with the field, the custom validation program being executed on a possible value.
14. The method as claimed in claim 13, wherein the custom validation program is a JavaScript program.
15. The method as claimed in any one of the preceding claims, wherein a field mask is associated with the field, the field mask checking that a possible value conforms with a predefined string pattern.
16. The method as claimed in any one of the preceding claims, wherein a possible value is derived from a selection list, or combination list, involving previously recognised responses.